

PROBLEMÁTICA DE LA EVALUACIÓN

Por ARSENIO PACIOS

1. *Sociedad y sistema educativo*

La correspondencia recíproca entre una sociedad en determinado grado de desarrollo y su sistema educativo es un hecho frecuentemente constatable, aunque tampoco falten ejemplos de desfase, bien por un anquilosamiento, bien por una evolución demasiado acelerada del último. En todo caso estos desfases constituyen un grave inconveniente tanto para el correcto desarrollo de las sociedades cuanto para la eficacia de los sistemas educativos inadecuados.

La conveniencia de esta correlación educación-sociedad es tan manifiesta porque en el fondo es la expresión de un derecho natural inherente a la persona humana: el de acceder a un nivel de educación en consonancia con el grado de desarrollo técnico, económico y cultural de la sociedad de que forma parte ¹.

Ahora bien, las sociedades de los distintos países evolucionan con muy distintos ritmos, que actualmente son con frecuencia extraordinariamente acelerados. Ello ha hecho que las reformas educativas se hayan convertido, en estos países, en un esfuerzo habitual y continuo, con el fin de ir adecuando sobre la marcha el tipo de sistema educativo adoptado al tipo

¹ JUAN XXIII, Encíclica *Pacem in terris*: "Nace también de la naturaleza humana el derecho a participar en los bienes de la cultura y, por tanto, a una instrucción fundamental y a una formación técnico profesional *en consonancia con el grado de desarrollo de la propia comunidad política*".

continuamente renovado de la sociedad. Lo que equivale a admitir que normalmente las reformas educativas van —o deben ir— pisándole los talones a las etapas de evolución de las sociedades a las que sirven. Si malo es que la educación se distancie del tipo de sociedad en que está inserta, por permanecer ininmóvil cuando debiera evolucionar, malo es también, y quizá peor, por arriesgado, el pretender establecer un sistema educativo demasiado avanzado pensando en una sociedad que sólo tendrá realización en el futuro, aun en el caso en que los cálculos hechos para preverla hayan de resultar acertados².

Por otra parte se desprende también de lo dicho que las reformas educativas deben proceder por evolución y desarrollo paulatinos y continuos y nunca por saltos ni por movimientos discontinuos, ya que la evolución social, aunque rápida, es también continua y sin solución de continuidad. Sólo en países que han dado en breves años un salto en su desarrollo pasando de un estado primitivo al designado hoy como de subdesarrollo pueden permitirse ensayos de reformas revolucionarias. En estos casos podrían justificarse tan enérgicas medidas sólo porque en dichos países hay que partir de cero y no se corre el riesgo de desechar fórmulas tradicionales que venían dando resultado para etapas anteriores de su evolución. Y aun así sería muy discutible la afirmación de que existen países en los que los administradores de la educación pueden partir de cero, ya que hasta en las sociedades más primitivas hay elementos educativos dignos de ser conservados. Pero lo

² Cuanto aquí vamos diciendo no debe ser entendido como una adecuación del sistema educativo a una sociedad actual en sentido absoluto. Es bien sabido que la educación debe tener a la vista las condiciones reales que presumiblemente ha de tener la sociedad en el momento en que los alumnos que están siendo formados hayan de incorporarse activamente a ella. Precisamente dentro de este contexto debe ser entendida y calibrada la prevalencia de la formación intelectual, moral, religiosa y política de los jóvenes sobre la mera instrucción, con lo que se confiere a la educación un sentido de anticipación que dispone al educando a hacer frente con garantías de éxito a los problemas, ocasionalmente inéditos, que en el futuro se le puedan presentar.

que parece claro es que ni los sistemas educativos son intercambiables en las distintas sociedades, ya que cada una de ellas tiene sus peculiares condiciones y su propia tradición educativa, ni se puede cometer la imprudencia, al acometer una reforma educativa, de proceder por el sistema del borrón y cuenta nueva, desconociendo el inmenso acervo de excelentes soluciones que la tradición ha acumulado y que merecen, no ser rechazadas sin más, sino adaptadas, mejoradas, sustituidas o completadas con prudencia y lucidez. De no hacerlo así provocaremos imprudentemente una disociación entre sociedad y educación, cuando el criterio para hacer evolucionar los sistemas educativos, que al fin y al cabo son medios, debe ser el grado y el sistema de desarrollo de la sociedad a la que deben servir. En suma una educación retrasada arrastra el retraso de la sociedad en su evolución; un sistema educativo que se adelante al nivel de desarrollo de una sociedad necesariamente causará trastornos en ella y se mostrará ineficaz y estéril. Y no faltarían ejemplos del presente y de la historia para fundamentar estas afirmaciones.

Todo lo que llevamos dicho sobre la necesaria coherencia entre el sistema educativo y una sociedad dada pone de manifiesto que las condiciones y las necesidades de una sociedad determinada son el criterio que sirve para fijar el tipo de sistema educativo que se ha de adoptar. Este criterio fundamental es extrínseco al sistema educativo como el fin lo es al instrumento. Pero existen además criterios intrínsecos al propio sistema como son los del control de los resultados del mismo, de que me propongo hablar. Estos criterios —exámenes, evaluación— confirman o desautorizan un sistema dado, pero atendiendo exclusivamente a su coherencia interna. Mas esto no basta para afirmar que en sí el sistema educativo sea válido. Sólo lo será si sirve eficazmente a la sociedad que lo utiliza. Y es que los métodos de control pertenecen al mismo sistema. Sin embargo, es necesario hallar sistemas de control que no sólo se adecúen al propio sistema educativo, sino que

al mismo tiempo pongan de manifiesto la adecuación de éste con la sociedad que lo utiliza.

2. *Los exámenes tradicionales y la sociedad actual*

Si estamos en lo cierto cuando exponemos las anteriores reflexiones, parece obvio también que los exámenes tradicionales, que en otro tiempo fueron fórmulas eficaces para calificar un sistema educativo coherente con el grado de desarrollo de aquellas sociedades, no responderán con la misma eficacia ante las condiciones actuales. En ese caso se hará preciso adoptar fórmulas de control y selección adecuadas a las necesidades presentes. A la evolución de la sociedad deberá seguir una evolución de los sistemas educativos y de sus procedimientos de control.

Parece que los significativos progresos tecnológicos, económicos y sociales de la sociedad tienen que influir sobre el contenido de los programas escolares y sobre todo el sistema educativo, incluidos los exámenes.

En primer término estos progresos han modificado la distribución de la población activa entre las distintas ramas de actividades. A lo largo de los últimos decenios se ha acelerado la disminución de empleos en el sector primario (agricultura, minería y pesca), se han ampliado los del sector secundario (industria) y se han multiplicado los del sector terciario (distribución, servicios, investigación, desarrollo). Todos estos cambios han tenido como consecuencia inmediata la explosión escolar en la enseñanza media, el acceso a este tipo de enseñanza de muchos alumnos pertenecientes a clases sociales anteriormente alejadas de ella, la renovación de la gama de materias estudiadas y la aparición entre los alumnos de nuevas ideas sobre sus carreras futuras. Ello obliga a replantear todo el sistema educativo y a investigar sobre estos temas con el fin de calificar los modelos aplicados o de encontrar otros nuevos más adecuados.

En efecto, el conjunto de los cambios que tienen lugar en los dominios social y de la enseñanza nos obliga a someter renovadamente a crítica la noción de enseñanza —sobre todo la noción de enseñanza media—, sus estructuras y su funcionamiento y, de un modo especial, los exámenes que venían siendo el criterio esencial de la perfección de la enseñanza. Y todo ello porque cualquier reforma de los programas y de los exámenes depende estrechamente de una valoración precisa de las consecuencias de una reorientación social³.

No parece difícil coincidir en que, en las naciones occidentales consideradas en conjunto, hasta 1945 se venían dando por supuestos principios como los siguientes :

- a) Los estudios medios son esencialmente propedéuticos para los universitarios.
- b) Se considera suficiente que entre un cinco y un diez por ciento de los alumnos en edad escolar estudien bachillerato.
- c) Carece de importancia que un 50 % o más de los alumnos de enseñanza media abandonen sus estudios en el camino, ya que con los restantes quedan suficientes para nutrir los cuadros de las actividades de gobierno y dirección del país.
- d) No debe existir preocupación por la suerte que puedan correr los fracasados.
- e) Por el contrario, la selección es esencialmente negativa, ya que tiene por misión descubrir los puntos flacos de los alumnos y no poner de relieve sus aptitudes diferenciadas.

A causa de esta actitud, ante el fracaso de los alumnos ni se estudiaban reformas de los programas, ni se trabajaba en

³ Véase el estudio de E. J. KING titulado "Etude sociologique et comparative des examens", en el folleto ciclostilado que con el título *Les examens de fin d'études secondaires* ha distribuido el Consejo de Europa (Comité de Coopération Cultural) en el otoño de 1971 (Doc. CCC/EGT (71)6).

mejorar los métodos pedagógicos, ni se creaban y utilizaban medios auxiliares nuevos, ni existía preocupación por adoptar dispositivos especiales para atender a los niños provinientes de medios rurales u obreros.

Los exámenes tradicionales pensados para un sistema educativo semejante no se adaptan evidentemente a las necesidades sociales actuales, aunque hayan sido instrumentos eficaces de selección en un estadio tecnológico y social menos avanzado.

A partir de la segunda guerra mundial (todos los países, aunque en distinto grado, presionados por las reivindicaciones sociales y políticas, se han preocupado de universalizar efectivamente la enseñanza primaria y de multiplicar la enseñanza media con ánimo de hacerla accesible al menos a todos los naturalmente aptos. Pero, en el fondo, esta nueva actitud, aunque aumentó el número de oportunidades ofrecidas, ni cambió la naturaleza de los programas y de los contenidos, ni introdujo modificaciones en el sistema de exámenes. En cambio el éxito se hacía cada vez más difícil para la multitud de candidatos al título de bachiller. Al mismo tiempo la excesiva cantidad de examinandos hacía problemática la normalidad de la selección de los mejores. La introducción de pruebas objetivas (que se pretendían inatacables por su carácter científico) como recurso de selección empeoró aún más las cosas.

En los últimos veinte años este panorama ha cambiado mucho, no tanto porque se hayan descubierto métodos perfectos de calificación y selección objetivas, sino porque se han puesto de relieve los defectos del sistema de evaluación y selección anterior. Hoy se sabe, por ejemplo, que pueden aplicarse calificaciones anormalmente bajas por muy variadas circunstancias y ajenas con frecuencia al propio proceso instructivo. Así los resultados de los exámenes pueden ser negativos simplemente porque las preguntas que se hacen al alumno versan sobre temas ajenos a su medio familiar y social, o porque éste carece de motivación suficiente por el simple hecho de pertenecer a medios sociales distintos de los del profe-

sor y de parte de sus compañeros. En efecto, el nivel de las notas obtenidas en un examen está en función de la experiencia escolar del candidato, pero también, y en alto grado, de su ambiente familiar y de sus aspiraciones personales. De aquí que, a pesar de la pretendida democratización, los alumnos procedentes de un ambiente "burgués" conservan sus ventajas tradicionales sobre los que proceden de un ambiente obrero. En parte explica este hecho el principio de que los alumnos obtienen éxito de acuerdo con las previsiones de sus profesores, porque los profesores tienden a identificar y premiar en sus alumnos las cualidades que encajan mejor con sus propias concepciones sociales⁴.

También desde el punto de vista cualitativo se están sometiendo a revisión nuestras actitudes ante la enseñanza. Hoy lo importante no es tanto los conocimientos tal como los conciben los profesores y examinadores cuanto el uso que se hace de ellos. Por un lado, la suma de los conocimientos que se estiman necesarios aumenta constantemente; por otro, la utilidad directa de los conocimientos adquiridos es puesta en duda a causa de la evolución de la sociedad en que aquellos han de ser aplicados. Cada vez resulta más indispensable no tanto acumular conocimientos cuanto hacer nacer nuevas aptitudes durante y después de la escolaridad. Al tiempo que asistimos a una transformación intrínseca de la naturaleza de los conocimientos y al aumento de su número y diferenciación, nos damos cuenta de que cada vez es más importante una buena formación con respecto a la simple información. Ahora bien la formación intelectual es tarea más ardua que la instrucción, y su evaluación por medio de los exámenes tradicionales presenta todavía mayores dificultades. En todas partes se está demostrando que estos exámenes no garantizan a los que los superan especiales éxitos ni en sus estudios ulteriores ni en las profesiones a que se dedican después.

⁴ Vid. E. J. KING en el artículo citado en la nota anterior.

Esto explica que Suecia, y con ella varios otros países europeos, hayan sustituido, paulatinamente y tras cuidadosas experiencias, su excelente sistema de exámenes de final del ciclo secundario por el sistema de evaluación continua.

3. *La evaluación como fórmula sustitutiva de los exámenes tradicionales*

El sistema de control del aprendizaje designado con el nombre de evaluación es demasiado complejo para que pueda ser encerrado en una breve definición. En efecto, el término evaluación es tan genérico que comprende cualquier actividad tendente a recoger información referente a la eficacia de nuestras operaciones y a sus resultados. Este principio o pronunciamiento crítico es habitual y corriente en el hombre en todos los dominios. Aplicado al sistema educativo y a la actividad instructiva debería recaer sobre el propio docente, sus procedimientos y su destreza operativa. Sin embargo, hasta el presente la evaluación en materia educativa viene recayendo casi exclusivamente sobre el alumno. Se juzgan sus actitudes, su madurez, sus adquisiciones, sus intereses y sus perspectivas de desarrollo. Pero sólo habrá evaluación completa si se tienen en cuenta también, y muy especialmente, los docentes, los métodos, la organización escolar y el ambiente de los alumnos. Por eso, aunque el término evaluación en su sentido amplio comprende los exámenes tradicionales de fin de ciclo, los exámenes periódicos mensuales o trimestrales o de fin de curso y cualquier otro medio de valoración de los resultados de la actividad instructivo-educativa, se aplica hoy con carácter casi privativo a un tipo de valoración más completo, comprensivo de todos los elementos integrantes de la relación educativa y de todos los aspectos de la actividad escolar que puedan ayudarnos a hacer un juicio exacto y completo de sus aciertos y sus errores.

Los exámenes tradicionales cubrían buena parte de los cometidos de la evaluación tal como aquí la estamos entendiendo; pero no todos ni con la finura y eficacia con que ésta pretende hacerlo. Los exámenes responden a una concepción *selectiva* de la finalidad de la escuela, lo que supone unos jueces que determinan los alumnos que deben ser aceptados o rechazados, sin tener en cuenta la parte de culpa o mérito que les corresponde a los profesores y a los centros educativos. Además estos exámenes pronuncian un juicio eminentemente negativo: libran a la sociedad o a los ulteriores niveles escolares de los alumnos ineptos, pero no se preocupan en modo alguno de señalar los procedimientos necesarios para colmar las lagunas y deficiencias encontradas en los alumnos y en el propio sistema educativo, con lo que dicho juicio resulta pedagógicamente estéril. Es una evaluación de carácter fiscal, pero que no intenta encontrar las causas y los remedios del fracaso, con lo que se malogra el efecto retroactivo que una correcta evaluación puede producir en la organización del proceso educativo con vistas a sacar el máximo provecho de las aptitudes de cada alumno.

Una auténtica evaluación escolar debería desempeñar al menos las siguientes funciones:

- a) Valoración general de la acción educativa realizada, a fin de hacer en ella las modificaciones necesarias en lo referente a contenidos, ritmo de aprendizaje y métodos.
- b) Valoración individual de las capacidades y las adquisiciones, de forma que permita la necesaria individualización del trabajo colectivo.
- c) Acopio progresivo de datos capaces de proporcionar material para una valoración de orden superior con fines de *orientación* y, en circunstancias determinadas de *selección* positiva y negativa.

Si bien estas tres funciones exigen fórmulas de valoración continua, de ello no se sigue una condenación absoluta de los exámenes de estado y de madurez. Lo que se deduce es la necesidad de una reforma urgente en el sentido de que el juicio se pronuncie no sólo fundado en un conjunto de pruebas, sino también en la atenta consideración de un historial académico personal que contenga toda una serie de evaluaciones continuas, que tendrán tanto más peso cuanto mayor sea la coincidencia de los juicios de docentes y examinadores⁵.

Dentro de la evaluación continua se pueden precisar todavía dos tipos: la *periódica* y la *acumulativa*, siendo la segunda la que en sentido propio recibe el nombre de verdadera evaluación continua. La periódica consiste en reemplazar el examen terminal único por un determinado número de pruebas a que los alumnos se someten en diferentes etapas de sus estudios (mensuales, trimestrales, anuales, etc.). Este tipo de evaluación venía siendo bastante común últimamente en varios países europeos y permite seguramente mejorar la validez de las pruebas al suprimir o atenuar algunos de los riesgos fisiológicos y psicológicos inherentes a los exámenes terminales únicos; pero no se salva de las críticas que se pueden formular legítimamente contra los exámenes escritos u orales tradicionales en cuanto instrumentos de medida. Al aumentar el número de las pruebas aumenta hasta cierto punto su validez y fiabilidad, pero se mantiene el riesgo de medir la facultad de retención más bien que la inteligencia o la creatividad.

La *evaluación continua acumulativa* se basa en la apreciación general del trabajo del alumno en cada materia durante toda su carrera. Este tipo de evaluación evita todos los inconvenientes de los exámenes tradicionales, pero abre el camino más fácilmente al juicio subjetivo, ya que el profesor suele ser el único juez en contacto con el alumno durante todo el ciclo y él sólo es el responsable de la evaluación.

⁵ ALDO VISALBERGHI, "Il problema della valutazione", *Bolletino del Ministero de la Publica Istruzione*, núm. 6, noviembre-diciembre 1971.

En la práctica y con el fin de evitar los inconvenientes de la evaluación periódica y de la continua se han utilizado ambas simultáneamente. Pero los resultados tampoco han sido suficientemente satisfactorios. Ello ha llevado a ensayar un tercer tipo de evaluación continua, que se podría denominar *evaluación global*, según propone Peterson⁶. Esta evaluación global contempla todo el conjunto de la personalidad y los antecedentes del alumno. Este método constituye indudablemente el mejor criterio de orientación y de selección, aunque tiene el inconveniente de ser mucho más subjetiva, y por lo tanto mucho menos segura, que la evaluación acumulativa.

En un sistema ideal de enseñanza no se necesitaría seleccionar los alumnos, sino sólo orientarles de acuerdo con sus aptitudes y los resultados escolares obtenidos. En ese caso la evaluación continua global sería un método excelente. Pero en la realidad las restricciones presupuestarias, la escasa demanda de universitarios en comparación con los presuntos candidatos, las necesidades del mercado de empleo, la falta de plazas universitarias, el sistema de concesión de becas limitadas, hace indispensable, quiérase o no, un sistema de selección de candidatos. Ello hace que los ensayos de orientación más serios que se han hecho en Europa se sitúen en el tránsito del bachillerato elemental al bachillerato superior y no en el paso del bachillerato superior a la universidad donde de un modo u otro aparece siempre cierto tipo de selección. Sin embargo, aun en aquel caso las técnicas precisas para hacer con garantías una evaluación eficaz de los alumnos hacen que gran parte de los profesores, aun los buenos, no pueda hacerse responsable de una orientación que puede condicionar toda la vida de un alumno. De ahí que en algunos casos —Francia es uno— se prevé la posibilidad de un examen supletorio cuando los padres realizan el consejo de orientación recibido por sus hijos, examen que puede contradecir los resultados de la evaluación

⁶ A. D. C. PETERSON, "Nouvelles techniques d'évaluation du travail scolaire", *Bulletin d'information du comité de l'Enseignement Général et Technique*, núm. 3, 1971. Consejo de Europa. Estrasburgo.

hecha en la escuela. Se ve bien cómo difícilmente se halla una orientación en estado puro, basada en la evaluación global. Por el contrario muy frecuentemente encubre una determinada selección y a veces desemboca simplemente en ella.

La responsabilidad que recae sobre la evaluación cuando de alguna manera desemboca en la selección de los alumnos plantea serios problemas referentes a su validez y fiabilidad. Otros problemas derivan de las técnicas y métodos que son precisos para llevarla a cabo. En fin nuevos problemas derivan de factores externos a la misma evaluación, pero que pueden deteriorarla y hacerla estéril o contraproducente. De esa problemática pasamos a ocuparnos a continuación.

4. *Los problemas de la evaluación*

Aunque muchos de los problemas que plantea la evaluación, buena parte de los cuales distan de estar resueltos, son comunes a cualquier clase de evaluación y a cualquier dominio de la realidad evaluable, otros son privativos de alguno de sus tipos. Así la evaluación de los resultados escolares o del aprovechamiento presenta menores dificultades que la de los elementos no cognitivos del proceso instructivo y de la personalidad de los alumnos. Al fin y al cabo en la evaluación de los conocimientos existe ya una larga experiencia, mientras que no acontece lo mismo en lo referente a factores afectivos, de interés, de aptitudes, de integración social, etc. Además estos dominios no cognitivos distan mucho de estar perfectamente definidos y ni siquiera influyen todos en el mismo grado en los resultados escolares.

En todo caso parece que se pueden distinguir dos grandes grupos de problemas de evaluación: los *intrínsecos*, que afectan directamente al mismo sistema, a sus características a los métodos y técnicas que se requieren, a las fuentes de datos y a su interpretación y utilización, y los *extrínsecos*, que derivan del personal docente y técnico que la ha de realizar, de

los centros, de los instrumentos, etc. En cada caso sería preciso considerar los que se puedan derivar del hecho de que se extienda o no la evaluación a los dominios no cognitivos del alumno.

Para mayor claridad quiero hacer algunas consideraciones previas sobre las dificultades que puede presentar la inclusión dentro de la evaluación de los factores no cognitivos, aunque estimo que tácticamente quizá convenga no acometer su evaluación hasta tanto la de los elementos propiamente cognitivos se haga con la suficiente perfección y seriedad como para permitirnos prescindir de los exámenes tradicionales que, quizá con alguna precipitación, han sido abolidos sin tener adecuadamente preparada su sustitución.

Me parece que se puede afirmar que el comportamiento humano está sujeto al influjo de factores cognitivos y no cognitivos. Aunque estrechamente relacionadas entre sí, estas dos categorías de factores pueden ser tratadas separadamente en función de su influencia sobre la educación. Se entiende por variables cognoscitivas aquellos aspectos de las capacidades intelectuales que han sido modificados por la experiencia y que, en el estado actual de nuestros conocimientos, son apreciadas por medios tradicionales o por técnicas más recientemente elaboradas, ampliamente utilizadas hoy en varios países europeos. Las variables no-cognitivas se refieren a los demás aspectos de la personalidad de los alumnos y a las características del medio ambiente que tienen influencia más o menos directa sobre los procesos de aprendizaje.

Parece indispensable disponer de evaluaciones objetivas en el dominio de lo no cognitivo cuando se trata de mejorar los procesos de aprendizaje o de prestar una ayuda efectiva al desarrollo personal. Hasta hoy estas variables no cognitivas en la educación han sido muy raras veces y muy accidentalmente tenidas en cuenta, bien para apreciar sus efectos sobre los procesos educativos, bien, a la inversa, para apreciar las modificaciones que los procesos educativos imprimen en las variables no cognitivas. Sin embargo, parece oportuno llamar la aten-

ción sobre la importancia de esas variables. Desgraciadamente hoy existen pocos instrumentos y métodos suficientemente fieles y válidos que ofrezcan un interés práctico en su evaluación, por lo que se hace preciso desarrollar investigaciones en este dominio y de una manera especial en los siguientes sectores:

- a) En función de los contenidos (madurez afectiva y social, sentido de las responsabilidades sociales, actitud ante el trabajo, sentido de los valores humanos, disposición para ayudar al prójimo, disposición para cooperar con los demás, aptitud para comunicarse, motivación ante el trabajo, interés por determinados campos, sentido comunitario).
- b) En función de la edad (aparición sucesiva de objetivos afectivos apropiados, correspondientes al desarrollo de las estructuras de la personalidad, de los valores humanos, de las emociones).
- c) En función de la personalidad (consideración de las necesidades individuales específicas durante la realización de los objetivos afectivos).
- d) En función de las instituciones y sistemas educativos (establecimiento de programas cuyos objetivos se hallen relacionados con las formas específicas de un sistema social dado).
- e) Formación del profesorado en métodos y técnicas de evaluación adecuadas.
- f) En función de estos mismos métodos, técnicas y procedimientos de valoración.

Quizá en el momento presente, y en espera de que sean elaborados nuevos métodos más eficaces, lo más urgente, sea acometer estudios e investigaciones en relación con las variables no cognitivas en dominios como el establecimiento de cri-

terios de evaluación, métodos, taxonomía de los rasgos del carácter, sistema de recogida y utilización de datos, etc.⁷

Me propongo ahora pasar revista a algunos de los principales problemas que puede plantear el decidido empeño de adoptar el sistema de evaluación en sustitución de los exámenes tradicionales. Aunque los he clasificado en intrínsecos y extrínsecos, no dejo de aceptar que algunos de ellos pertenecen a ambos dominios. Tampoco creo que se pueda justificar el orden en que los voy a exponer, ya que son varios los criterios de ordenación que legítimamente se podrían emplear. Estimo suficiente presentar una serie de interrogantes, problemas y dificultades que puedan incitar a los estudiosos a investigar en los distintos aspectos de tema tan complejo como es la evaluación.

1. *Fiabilidad y validez*. — Quizá sea este el problema más de fondo que se puede ofrecer. Si los exámenes tradicionales son rechazados porque muy a menudo dejan mucho que desear en estos aspectos, es de esperar que la evaluación los supla en ellos con ventaja. Si bien la evaluación continua (periódica, acumulativa o global) ofrece ventajas evidentes en lo tocante al efecto retroactivo y la rapidez, presenta por el contrario graves dificultades por lo que hace a la fiabilidad. Los métodos empleados para hacer la evaluación más segura y, sobre todo, para mejorar la comparabilidad entre los resultados de los distintos centros docentes presenta un interés considerable. En Suecia han llegado laboriosamente a la construcción de pruebas objetivas que permiten calcular el nivel nacional medio de conocimientos para cada curso de enseñanza secundaria superior. De acuerdo con este nivel han construido pruebas normalizadas que aplicadas a los alumnos en todo el país sirven de indicadores a los profesores para apreciar si el nivel general de su clase y las dispersiones de los alumnos están conformes con la media nacional. Sin

⁷ Para lo referente a las variables no cognitivas ha tenido a la vista el documento provisional elaborado por el Simposio del Consejo de Europa celebrado en Berlín del 12 al 19 de noviembre de 1971. También he tenido a la vista este documento para otros temas del presente trabajo.

embargo, sigue habiendo problemas sin resolver. Uno se imagina bien el resentimiento que puede producir el hecho de que pequeñas diferencias entre los resultados de las pruebas normalizadas y la calificación otorgada por el profesor decide el ingreso o no del alumno en el centro superior al que desea acceder. Una cosa parece cierta: la aplicación del sistema de evaluación, que parece eficaz para la obtención de títulos o para la orientación, no es instrumento suficientemente fino y eficaz para la selección de los alumnos, a no ser que los profesores sean especialmente formados para la aplicación de dicho sistema. Otro peligro acecha a la evaluación: no sería infundado temer que, si se tiene en cuenta la formación requerida, el tiempo empleado en reuniones y la amplitud de los trabajos de evaluación, estas técnicas complejas, aplicadas por los profesores, son demasiado costosas en relación con los resultados que proporcionan y que el interés y el entusiasmo que despiertan en el período de experimentación pueden desaparecer desde el momento en que se generalice su empleo.

En cuanto a la validez de la evaluación continua parece ser que supera notablemente a la del sistema de exámenes tradicionales. Así, mientras que en el sistema tradicional de pruebas la calificación suele resultar falsedad por el hecho de versar sobre los resultados de un alumno en una ocasión dada influida por las llamadas "condiciones de examen", la evaluación continua se extiende a lo largo de todo el proceso instructivo, contempla todos o la mayor parte de los objetivos propuestos y se realiza en "condiciones normales" para el escolar. Con frecuencia los exámenes no miden las aptitudes y los conocimientos que deben medir, y cuando lo hacen esos conocimientos y aptitudes no son verdaderamente los que les han de servir en el futuro. También se les ha reprochado los nefastos efectos retroactivos, al impulsar a los alumnos a aceptar las ideas y las opiniones del profesor o del libro en vez de ejercitar su inteligencia y su espíritu de crítica. Pero antes de condenar categóricamente los exámenes y adoptar una forma de evaluación más continua, habría que preguntarse por qué han gozado tanto tiempo de favor en

los sistemas educativos. Entre otras razones se podría aducir que esas cualidades y aptitudes que ponen de manifiesto los exámenes son muy importantes y deben ser tenidas en cuenta por cualquier evaluación. Por otra parte, al menos para muchos alumnos, los exámenes son altamente estimulantes, competitivos, y les ofrecen la ocasión de preparar una materia con una finalidad precisa, ordenada y sistemáticamente. Por fin, la verdadera razón por la que los exámenes siguen desempeñando un importante papel en nuestros sistemas de calificación, sobre todo en los momentos decisivos del proceso educativo, es sin duda su fiabilidad y su imparcialidad manifiesta. La imparcialidad, que puede ser demostrada a los padres y a los alumnos, es de capital importancia siempre que el porvenir de un individuo depende de su éxito o de su fracaso en un examen. En cuanto a la fiabilidad, aunque débil, es sin embargo superior a la mayor parte de las técnicas de evaluación continua. Si queremos reemplazar los exámenes de este tipo por un método de evaluación continua que no tenga sus defectos intrínsecos, habrá que esforzarse por lograr que la nueva fórmula conserve el mayor número posible de las virtudes de la antigua ⁸.

2. *Objetividad*. — A pesar de las críticas que desde el punto de vista de la objetividad se han hecho a los exámenes tradicionales, no parece que se haya logrado aumentar esta cualidad del juicio con la aplicación de la evaluación continua. El hecho de que sea el profesor único juez condiciona seriamente la objetividad de su juicio. Ciertamente que el empleo masivo de pruebas objetivas, que cubran prácticamente todo el contenido de los programas, puede introducir un apreciable factor de objetividad en la evaluación de los alumnos. Pero las pruebas objetivas sólo pueden constituir una parte de las fuentes de obtención de datos y, además, estas pruebas que son un excelente medio de medida, deben ser elaboradas, interpretadas y califi-

⁸ A. D. C. PETERSON, *ob. cit.*

cadadas de acuerdo con baremos siempre discutibles y que dependen en gran parte de las condiciones personales del profesor.

3. *Valor autorregulador de la evaluación.* — La evaluación continua supera ampliamente a los exámenes tradicionales en su valor para establecer sobre la marcha las correcciones precisas en los contenidos, los métodos y el ritmo de la enseñanza, así como sobre todo el complejo sistema educativo. Pero también es cierto que las pruebas propuestas sólo serán capaces de servir de criterios correctores del sistema cuando no sólo se limiten a detectar el grado de aprovechamiento del alumno, sino cuando, además, sean capaces de determinar el camino seguido por aquél en su aprendizaje, ya que esto puede tener más influencia en el estudio de nuevos y mejores métodos, adaptados a la individualidad de cada alumno, que el simple hecho de calificar equitativamente unos resultados. Así, dos alumnos que hayan cometido el mismo número de faltas en un dictado no solamente pueden diferenciarse entre sí por la naturaleza y la calidad de las faltas, sino —y esto es más interesante— por el tipo de proceso, distinto en cada uno, que les ha conducido a las mismas faltas. En unos casos podrá ser por poco ejercicio de lectura, en otros por ignorancia de las reglas de la ortografía, en otros porque el alumno procede de un medio cultural muy bajo, en otros porque procede de una región en que existen defectos generales de pronunciación, etc. Y en cada caso la terapéutica a aplicar será distinta. Ello nos hace ver cuán difícil ha de ser construir pruebas que, lejos de limitarse a detectar el grado de conocimiento del alumno, procuren además poner de manifiesto el camino o proceso seguido por la mente hasta expresar la contestación a la cuestión propuesta. Parece evidente que sin la ayuda de personal técnico muy especializado no serán capaces los profesores de aprovechar estas ventajas de la evaluación.

4. *Análisis de objetivos.* — La validez de las pruebas de la evaluación supone el previo conocimiento y análisis de los objetivos que se propone cubrir la enseñanza. Una prueba es

válida si mide lo que se quiere medir. Para ello hay que conocer los objetivos no sólo globalmente, sino analíticamente. Y pocos profesores están entrenados en distinguir, clasificar y ordenar jerárquicamente los objetivos de su materia. Parece que su labor se vería grandemente facilitada si estuviesen entrenados en la programación de su enseñanza de acuerdo con el método de la enseñanza programada. Pero esta preparación son muy pocos los que la poseen. En todo caso creo que se les podría ayudar mucho si se construyen, a nivel nacional, y para cada materia y curso, un abanico de objetivos prefijados, debidamente estructurados, señalando en cada caso cuáles son los imprescindibles, los muy convenientes y los accesorios. Pero es un trabajo éste que está totalmente por hacer.

5. *Métodos y técnicas.* — La evaluación objetiva de los resultados del proceso instructivo, aunque muy necesaria, sólo muy parcialmente se realiza en la actualidad. Con frecuencia los profesores aprecian todavía los hechos de una manera intuitiva, a falta del dominio de las técnicas necesarias para observar, proponer pruebas idóneas de medida e interpretar los datos obtenidos. Es preciso proporcionar a los profesores un mínimo de conocimiento sobre los elementos básicos de la medida y de la teoría de los tests. El profesor debe saber que la apreciación de fenómenos complejos exige un análisis de los diferentes componentes que los integran. Los instrumentos utilizados (cuestionarios, tests, etc.) deben estar siempre concebidos en función del uso previsto y de la información que pueden dar. Al lado de los métodos tradicionales, se debería utilizar un amplio repertorio de técnicas de evaluación, como los métodos de observación, de entrevista, de apreciación personal. Asimismo se habrán de tener en cuenta las aptitudes y las experiencias anteriores del alumno para adaptar la instrucción a diversos grupos de alumnos, de profesores y de material de enseñanza, de forma que los métodos de evaluación contribuyan también al proceso educativo considerado en su conjunto.

6. *Fuentes de datos.* — De la correcta dosificación de las fuentes de donde se han de tomar los datos para la evaluación dependerá en buena parte la justeza y calidad de ésta. Las pruebas objetivas y tests bien escogidos permiten conferir a la evaluación un alto valor de objetividad. Pero una evaluación fundada exclusiva o preponderantemente en estos medios sería siempre incompleta e inexacta. Será pues necesario apelar a otras fuentes de datos como son la observación de la actividad del alumno, las preguntas y respuestas verbales en clase, los coloquios o diálogos y, sobre todo, los ejercicios escritos frecuentes sobre temas variados. Para que la objetividad no se resienta con la aplicación de estos medios habrán de aplicarse técnicas de análisis y descomposición de estos procesos complejos a fin de mejorar la exactitud de su medida.

7. *Utilización de los resultados de la evaluación.* — El uso de estos resultados presenta aspectos deontológicos que no sería lícito olvidar. Parece que es esencial controlar severamente el uso de las informaciones obtenidas por los procedimientos de evaluación, especialmente en el dominio no cognitivo. Puesto que la razón de ser de estos datos es la mejora del proceso educativo, deberían ser conservados dentro del sistema escolar y tratados como estrictamente confidenciales. Incluso dentro de la misma institución escolar, una difusión incontrolada de estas informaciones supondría peligros, por lo que los resultados de la evaluación no deberían ser comunicados más que a las personas directamente interesadas.

Sin embargo, de hecho, todo está sistemáticamente establecido en contra de estas restricciones al uso de la información obtenida. Los centros superiores, antes de recibir a determinados alumnos, exigen saber cuanto sea posible de sus estudios y experiencias anteriores; el tránsito de unos centros a otros presenta las mismas exigencias; la concesión de becas y bolsas de estudio se condiciona también a méritos que son reflejo fiel de los resultados de la evaluación. Todo ello plantea problemas

muy concretos acerca de la licitud, en cada caso concreto; de la utilización de las informaciones obtenidas en la evaluación.

8. *Personal docente y técnico.* — Avanzando más hacia los problemas que hemos dado en considerar extrínsecos a la evaluación parece que debemos fijar nuestra atención sobre los problemas que se derivan de la exigencia de que el personal docente y técnico que en ella interviene tenga la preparación necesaria.

Se hace preciso lograr una estrecha colaboración entre los docentes y los especialistas, como psicólogos y consejeros de orientación, si es que se aspira a desarrollar nuevos procedimientos de evaluación, aplicarlos y explotar los resultados obtenidos. Ello es especialmente indispensable si se pretende integrar la orientación continua en el proceso educativo total.

El tema de la evaluación deberá figurar siempre en los cursos de formación del profesorado, tanto a lo largo de su formación inicial como en el de su perfeccionamiento posterior. En la *formación inicial* debería prestarse especial atención a las técnicas de desarrollo del plan de estudios de cada materia (especificación de objetivos, determinación del orden de presentación, medida de los resultados); a los problemas generales de la evaluación, como son los fines, funciones y métodos de la misma (docimología, principios de la construcción de tests); y al nuevo papel del profesor en una enseñanza más individualizada. El *perfeccionamiento posterior* debería versar, entre otras cosas, sobre la construcción y experimentación de diversas técnicas de evaluación, de pruebas de conocimiento, construidas por los mismos docentes para sus particulares necesidades. Aquellos que tienen que experimentar y desarrollar nuevos procedimientos de evaluación verían considerablemente facilitado su trabajo si dispusieran de medios técnicos modernos como equipos de reproducción, medios audiovisuales y centros de informática, que les permitiesen, por ejemplo, la constitución de bancos de datos o de ítems y un análisis más detallado del proceso de enseñanza y aprendizaje.

9. *La diversidad de centros docentes.* — Dadas las precarias condiciones de fiabilidad, validez y objetividad de la evaluación continua, nuevos y graves problemas se derivan, no sólo de la diversidad de evaluadores —cada profesor es uno y diferente de los demás—, sino sobre todo de las diferentes clases y categorías de centros docentes, dominados todos por una tradición y unos intereses tan variados.

En cuanto a los docentes puede paliarse en parte la disparidad de criterios recurriendo a la evaluación en equipo. Sin embargo, serían de esperar mejores resultados si la asociación en equipo no se realizara sólo en el momento de la evaluación de grupos determinados de alumnos, constituyendo el llamado *team teaching*.

Por lo que hace a la diversidad de centros, el problema se ve agravado considerablemente dado el gran peso que en España tiene la enseñanza no oficial. Es éste un problema de capital importancia, que puede comprometer muy seriamente el éxito del sistema de evaluación recientemente implantado. Parece que un remedio válido sería una correcta formación del profesorado en estos dominios, no sólo técnica, sino también deontológica. Una correcta organización interna de los centros ayudaría probablemente a la resolución de este problema.

10. *La aceptación por parte de los alumnos y sus padres.* Ya hemos visto cómo los resultados de los exámenes tradicionales gozan de la confianza y el asentimiento de padres y alumnos, aun cuando sean muy numerosas también las censuras que contra ellos se mueven. No acontece lo mismo con la evaluación, que no puede exhibir en el mismo grado la objetividad con que se ha realizado. Esto y la resistencia a la novedad hace que se ofrezcan reparos a la introducción del nuevo sistema. Parece que un buen medio para contrarrestar esta actitud es la práctica de la autovaloración por parte de los alumnos, no en el sentido de que ellos mismos se autocalifiquen, sino ofreciéndoles la oportunidad de discutir los resultados de las pruebas objetivas. Con esta práctica se obtienen muchas ventajas. En primer lugar los

alumnos se hacen conscientes de su propio valor real, al menos en las áreas cubiertas por dichas pruebas, lo que constituye un excelente punto de partida para una futura autoorientación; en segundo lugar, de esta discusión puede derivarse el análisis y la consiguiente corrección y perfeccionamiento de las propias pruebas, al descubrirse el sentido equívoco de muchas preguntas, la falta de análisis y la confusión de otras, e incluso el rechazo de tipos de preguntas poco significativas y hasta inexactas.

En todo caso parece que las calificaciones por evaluación tardarán bastante en ganarse la confianza de padres y alumnos, lo que exige un esfuerzo suplementario para acrecentar rápidamente su prestigio a base de seriedad, honestidad y eficacia a la hora de aplicarlas desde el primer momento.

Otros problemas que afectan directamente a padres y alumnos, se derivan de las nuevas notaciones adoptadas. Acostumbrados a los valores numéricos, extraordinariamente aptos para discernir diferencias mínimas entre alumnos a la hora de pretender becas u otros beneficios similares, se encuentran con que las recientemente adoptadas son instrumento demasiado grosero para llegar a tales discriminaciones. En este sentido resulta aleccionador que las familias de los alumnos de enseñanza secundaria de Bélgica se resisten a aceptar las nuevas calificaciones que, para ellas, carecen de sentido o resultan escasamente significativas.

Algunos de los problemas más típicos de la evaluación continua en relación con la situación española propiamente dicha han sido objeto de reflexión especial en un artículo publicado recientemente (año 1971) en la revista *Organización Escolar* bajo el título "La evaluación continua del alumno y su problemática".

5. *Lo que habrá que hacer*

Aunque al hilo de las reflexiones precedentes ya se han ido señalando gran parte de las tareas que nos esperan para

perfeccionar y adecuar el instrumento de calificación que es la evaluación, creo que no resultará ocioso llamar la atención expresamente sobre lo que resulta urgente realizar si queremos disponer de un medio de evaluación seguro y completo del proceso educativo de los alumnos.

Quizá haya que dar una prioridad absoluta a la labor investigadora, que creemos debe cubrir objetivos muy variados. Entre otros se hallan los referentes a una definición lo más exacta posible del contenido y las principales variables tanto del dominio cognoscitivo como del no-cognoscitivo, campo este último que necesita aún una amplia y profunda elaboración. También parece indispensable la promoción de investigaciones referentes a la evaluación como medio de selección de alumnos, de orientación, de otorgamiento de títulos y de diplomas, de concesión de becas y becas. Cada una de estas finalidades presenta características y exigencias especiales que parece no pueden ser cubiertas todas ellas indiscriminadamente por la evaluación genérica tal como la concebimos ahora.

Será necesario también descubrir y construir nuevos métodos e instrumentos de evaluación para ponerles a disposición de los profesores. A este fin se necesita la creación de un organismo técnico central que se ocupe del análisis y catalogación de objetivos por cada materia, de la construcción de pruebas objetivas numerosas y variadas, de la redacción de instrucciones inmediatamente aplicables a la práctica, de la resolución de problemas de valoración de pruebas, si es preciso con el empleo de computadores, y del asesoramiento a los centros y profesores en estas materias.

También habrá que estudiar primero —para luego divulgar los resultados de estos estudios— todo lo referente a las distintas fuentes posibles de datos para la evaluación, teniendo en cuenta las distintas finalidades que se deben alcanzar: la determinación de los resultados de la instrucción y formación, el rendimiento del alumno supuesto un determinado nivel de aptitudes, las distintas variables de su personalidad, incluidas las afectivas y ambientales, la estimación global de su evolución

educativa determinando sus principales elementos integrantes y el modo como deben ser tenidas en cuenta a la hora de dar expresión a esta apreciación.

Sin embargo, sigue siendo uno de los principales problemas merecedores de nuestra atención el hecho mismo de la realización de la evaluación. Aunque desde el punto de vista teórico estuvieran resueltos todos los problemas, seguiría siendo lo más difícil su puesta en práctica. Para lograr esto adecuadamente parece imprescindible concentrar nuestra atención urgentemente en la formación del profesorado en las distintas técnicas de evaluación y en el empleo metódico y racional de las diversas fuentes de los datos que han de servir de base para un juicio exacto, completo y equilibrado de los alumnos. Esta formación debería ser a mi juicio, sobre todo en el principio, eminentemente práctica.

Mientras todo esto no se logre la evaluación será un empeño difícil y de resultados mediocres, poco válido y casi nada fiable. En este caso sería cosa de pensar si no sería bueno conservar transitoriamente, además de esta evaluación insegura y balbuciente, alguno de los tipos de evaluación tradicionales mejorado y puesto al día, como, por ejemplo, la fórmula del examen-balance que de tanto favor —al menos en teoría— ha gozado en Francia los últimos años.